

# IMPACT OF LAND MODEL PHYSICS ON SOIL MOISTURE ASSIMILATION

*R. H. Reichle<sup>1</sup>, Sujay V. Kumar<sup>1</sup>, R. D. Koster<sup>1</sup>, C. D. Peters-Lidard<sup>1</sup>, and W. T. Crow<sup>2</sup>*

<sup>1</sup>NASA Goddard Space Flight Center, Greenbelt, MD

<sup>2</sup>Hydrology and Remote Sensing Lab., USDA/ARS, Beltsville, MD

## ABSTRACT

The skill of root zone soil moisture estimates derived from assimilating satellite retrievals of surface soil moisture into a land surface model depends, among other factors, on the specific structure of the land model. Based on a suite of synthetic data assimilation experiments with four different land surface models we find that it is prudent to use a simpler land surface model as part of the data assimilation system.

**Index Terms**— Soil Moisture, Data Assimilation

## 1. INTRODUCTION

A common approach to estimating soil moisture is to drive a land surface model (LSM) with observed meteorological forcing. The physical formulations within the LSM integrate the forcing data and produce estimates of soil moisture and associated land surface fields. These model products, however, are subject to error due to uncertainties in the meteorological forcing, faulty estimates of relevant parameters, and deficient LSM formulations. Indirect measurements (retrievals) of surface soil moisture can be obtained from satellite sensors that measure microwaves emitted by the land surface. The data coverage, however, is incomplete in space and time, and data are subject to measurement uncertainties and errors in the retrieval process.

Data assimilation systems are designed to merge the retrieval information with the spatially and temporally complete information provided by the LSM and produce a superior product (e.g. root zone soil moisture). The assimilation system acts to propagate the surface retrieval information into deeper soil layers, giving the retrievals an otherwise unobtainable relevance to such applications as the initialization of weather and seasonal climate forecasts. Land surface models, however, differ significantly in their representation of subsurface soil moisture processes and as a result, differ in how they propagate information from the surface into the deeper soil layers. In this paper, we investigate how the specific formulation of the land surface model impacts the information contribution of surface soil

moisture retrievals to root zone soil moisture products in a data assimilation system.

## 2. METHODOLOGY

The study is conducted using the recently developed Land Information System (LIS) data assimilation testbed, an interoperable framework for sequential data assimilation that enables the integrated use of multiple LSM's, multiple observations, and multiple data assimilation algorithms. We conducted a suite of Observing System Simulation Experiments (OSSE's) in which synthetic retrievals of surface soil moisture were assimilated into the Catchment, MOSAIC, Noah, and CLM models with the Ensemble Kalman Filter (EnKF). The experiment domain approximately covers the Continental U.S. for the period from 2001 to 2006. The LSM's vary in complexity in the representation of subsurface moisture dynamics. The Catchment LSM essentially describes deviations from the equilibrium soil moisture profile and has a relatively strong vertical coupling between surface and root zone soil moisture. By contrast, the layer-based models MOSAIC (3 layers), Noah (4 layers), and CLM (10 layers) have successively weaker coupling between their surface layers and root zone soil moisture (defined here as the soil moisture in the top 1 m of the soil column).

The OSSE consists of four steps: (1) A 12-member ensemble integration was performed for each of the four LSM's by adding perturbations to the surface meteorological forcing data and selected state variables. For each LSM, a single ensemble member was selected to represent the "true" land surface fields and the ensemble mean is used as the "open loop" (no data assimilation) estimate for that LSM. (2) Four corresponding sets of synthetic surface soil moisture retrievals were generated by adding suitable synthetic observation errors to the LSM true fields. (3) Each of the four synthetic retrieval datasets was assimilated into each of the four LSM's for a total of 16 data assimilation experiments (using the same perturbation settings as in the open loop ensemble integration). (4) The root zone soil moisture products from the open loop and the data assimilation integrations were then evaluated against the corresponding truth data. The experiments illustrate the

sensitivity of model parameterizations and physical representations on the efficiency of soil moisture assimilation.

### 3. RESULTS

Table 1 summarizes the results of the experiments. Listed in Table 1 is the normalized information contribution to the root zone soil moisture product for each of the 16 assimilation integrations. To compute the normalized information contribution, we first compute the anomaly time series correlation coefficient  $R_a$  of the root zone soil moisture assimilation product and the synthetic “truth” data. Next, we compute the same for the open loop (no assimilation) product ( $R_o$ ). The normalized information contribution is then defined as  $NIC = (R_a - R_o) / (1 - R_o)$  and measures how much of the maximum possible skill improvement ( $1 - R_o$ ) is realized through data assimilation ( $R_a - R_o$ ). Under the assumption that  $R_a \geq R_o$ , that is, the assimilation product is no worse than the model-only product, we have  $0 \leq NIC \leq 1$ . For  $NIC=0$ , the assimilation of surface soil moisture retrievals does not add information to the root zone soil moisture product. For  $NIC=1$ , the assimilation of surface soil moisture retrievals realizes the maximum possible skill contribution. The motivation for the normalization is that it is extremely difficult, if not impossible to achieve identical open loop skills for the different LSM’s across all 16 experiments.

Table 1 indicates two main results. By comparing the columns of Table 1 we can assess the potential of surface soil moisture assimilation under a range of possible subsurface physics. If the Catchment and MOSAIC LSM’s with their strong surface-root zone vertical coupling (not shown) represent nature well, improvements through assimilation of surface information yields on average the strongest improvements in root zone products ( $NIC=0.45 \dots 0.47$ ). If, on the other hand, models with weaker surface-root zone coupling such as Noah or CLM represent the true subsurface physics best, potential improvements through assimilation of surface observations are more limited ( $NIC=0.28 \dots 0.36$ ).

The second main result concerns the optimal choice of LSM for data assimilation under the assumption that the true subsurface physics are unknown. By comparing the rows of Table 1, we can see that choosing the Catchment and MOSAIC LSM yields on average higher skill improvements ( $NIC=0.47 \dots 0.50$ ) than choosing Noah ( $NIC=0.37$ ) or CLM ( $NIC=0.22$ ). This result is again compatible with the strength of the surface-root zone coupling and the complexity of the LSM. Unless it is clear that subsurface moisture dynamics are best represented by a complex LSM, it is therefore prudent to choose a simpler LSM in the data assimilation system.

<i>Data assimilated into:</i>	<i>Synthetic retrievals generated with:</i>				
	Catchment	MOSAIC	Noah	CLM	<i>(average)</i>
Catchment	0.71	0.54	0.36	0.38	<i>0.50</i>
MOSAIC	0.55	0.69	0.31	0.33	<i>0.47</i>
Noah	0.43	0.43	0.36	0.26	<i>0.37</i>
CLM	0.11	0.21	0.10	0.45	<i>0.22</i>
<i>(average)</i>	<i>0.45</i>	<i>0.47</i>	<i>0.28</i>	<i>0.36</i>	

Table 1: Normalized information contribution of assimilated surface soil moisture observations to skill in root zone anomalies. Columns indicate which model was used in the generation of the synthetic retrievals. Rows indicate which model was used to assimilate the synthetic retrievals. Last row and last column indicate averages across all models.